

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2003, 16:11:58 ; Search time 8.22857 seconds
(without alignments)
91.441 Million cell updates/sec

Title: US-09-905-691-5

Perfect score: 16

Sequence: 1 CRRARARARRARAE 16

Scoring table: OLIGO

Gapop 60.0 , Gapext 50.0

Searched: 127863 seqs, 47026705 residues

Word size: 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SwissProt_41.4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	43.8	263	CCMC_BRAJA	P30962 bradyrhizob
2	7	43.8	291	FM32_MYCTU	Q10515 mycobacteri
3	7	43.8	417	PROA_MEIRU	O86053 meiothermus
4	7	43.8	423	CES5_HUMAN	O9BXW7 homo sapien
5	7	43.8	428	Y486_MYCLE	P54138 mycobacteri
6	7	43.8	444	VGLX_HSVBS	O08103 bovine herp
7	7	43.8	474	CYAE_BORPE	P11092 bordetella
8	7	43.8	480	Y486_MYCTU	Q11152 mycobacteri
9	7	43.8	521	EX7L_RHTLO	O987V3 rhizobium 1
10	7	43.8	603	US26_HCNVA	P09699 human cytom
11	7	43.8	621	HEM1_AGABI	O92403 agaricus b1
12	7	43.8	742	UL47_HSVBP	P30021 bovine herp
13	7	43.8	777	METE_CAUCR	Q9AAW1 caulobacter
14	7	43.8	913	VGLB_PPRVF	P08355 pseudorabie
15	7	43.8	1318	VP14_EBV	P03179 Epstein-bar
16	7	43.8	1394	HAP_HAEIN	P45387 haemophilus
17	7	43.8	2390	SPCP_HUMAN	O15020 homo sapien
18	7	43.8	4523	DYHB_HUMAN	O96dt5 homo sapien
19	6	37.5	68	RPOZ_NEIMA	Q9Jq99 neisseria m
20	6	37.5	88	RPOZ_HAEIN	P43740 haemophilus
21	6	37.5	89	RPOZ_PASMO	Q9cmh2 pasteurella
22	6	37.5	90	RPOZ_VIBCH	O9kmh3 vibrio chol
23	6	37.5	90	RPOZ_VIBPA	O87tb0 vibrio para
24	6	37.5	90	RPOZ_VIBVU	O8ddv5 vibrio vuln
25	6	37.5	91	RPOZ_ECOLI	P08374 escherichia
26	6	37.5	91	RPOZ_YERPE	O8zjg3 yersinia pe
27	6	37.5	94	R28A_MYCTU	Q10879 mycobacteri
28	6	37.5	108	N1FW_RHOSH	Q01182 rhodobacter
29	6	37.5	113	UL67_HCNVA	P16747 human cytom
30	6	37.5	117	RL18_HAEIN	P44356 haemophilus
31	6	37.5	117	VGLJ_HSVSB	P36344 simian herp
32	6	37.5	120	PAND_ALCEU	O9zh15 alcaligenes
33	6	37.5	120	PAND_RALSO	O8xvu6 ralstonia s

RESULT 1

ID	CCMC_BRAJA	STANDARD	PRT	263 AA
AC	P30962			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Heme exporter protein C (Cytochrome c-type biogenesis protein cycz).			
GN	CYCZ OR CCMC OR BLR0469.			
OS	Bradyrhizobium japonicum.			
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;			
OC	Bradyrhizobiaceae; Bradyrhizobium.			
NCBI_TaxID=375;				
[1]				
RP	SEQUENCE FROM N.A.			
RC	STRAIN-1101R15;			
RX	MEDLINE-91210304; PubMed-1850420;			
RA	Ramseller T.M., Winteler H.V., Hennecke H.;			
RT	"Discovery and sequence analysis of bacterial genes involved in the			
RT	biogenesis of c-type cytochromes.";			
RL	J. Biol. Chem. 266:7793-7803(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-USDA 110;			
RX	MEDLINE-22484998; PubMed-12597275;			
RA	Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,			
RA	Sasamoto S., Watanabe A., Idesawa K., Iriuchi M., Kawashima K.,			
RA	Kohara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,			
RA	Tabata S.;			
RT	"Complete genomic sequence of nitrogen-fixing symbiotic bacterium			
RT	Bradyrhizobium japonicum USDA110.";			
RL	DNA Res. 9:189-197(2002).			
CC	-1- FUNCTION: REQUIRED FOR THE EXPORT OF HEME TO THE PERIPLASM FOR THE			
CC	BIODEGRADATION OF C-TYPE CYTOCHROMES.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane			
CC	(Probable).			
CC	-1- SIMILARITY: BELONGS TO THE CCMC/CYCZ/HELC FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M60874; AAA26194.1; ..			
DR	EMBL; AP005936; BAC45734.1; ..			
DR	PIR; C39741; C39741.			
DR	InterPro; IPR002541; CytC_asm.			
DR	InterPro; IPR003557; CytC_biol_Ccmc.			
DR	Pfam; PF01578; CytC_asm; 1.			
DR	PRINTS; PRO1386; CCMCBIOGNIS.			
DR	TIGRFAMs; TIGR01191; ccmc; 1.			
KW	Cytochrome c-type biogenesis; Transport; Transmembrane;			
KW	Inner membrane; Complete proteome.			

P49926 canis famil
O24709 synechococc
Q9hve9 pseudomonas
P52767 pinus thunb
Q07178 rhodobacter
O53412 mycobacteri
Q33060 mycobacteri
Q10542 mycobacteri
Q9yf92 aeropyrum p
Q17606 caenorhabdi
O83817 treponema p
P49694 thermomonos

ALIGNMENTS

34	6	37.5	122	1	CRF_CANFA
35	6	37.5	130	1	RS11_SYNPE
36	6	37.5	130	1	RS8_PSEAE
37	6	37.5	134	1	RK16_PINTH
38	6	37.5	135	1	NIU1_RHOCA
39	6	37.5	139	1	LPOV_MYCTU
40	6	37.5	143	1	Y880_MYCLE
41	6	37.5	143	1	Y880_MYCTU
42	6	37.5	145	1	RL32_ASRPE
43	6	37.5	162	1	RL24_CAEL
44	6	37.5	177	1	Y845_TREPA
45	6	37.5	180	1	YPRW_THUCU

FT TRANSMEM 19 39 POTENTIAL.
 FT TRANSMEM 61 81 POTENTIAL.
 FT TRANSMEM 92 112 POTENTIAL.
 FT TRANSMEM 126 146 POTENTIAL.
 FT TRANSMEM 157 177 POTENTIAL.
 FT TRANSMEM 198 218 POTENTIAL.
 SQ SEQUENCE 263 AA; 28831 MW; A02EF7576994EC0 CRC64;
 Query Match 43.8%; Score 7; DB 1; Length 263;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RAARAAA 9
 DB 151 RAARAAA 157

RESULT 2
 YH32_MYCTU STANDARD; PRT; 291 AA.
 ID YH32_MYCTU
 AC Q10515; Q10516;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein RV2232/RV2233.
 GN RV2232/RV2233 OR MT2292 OR MTCY427.13/MTCY427.14.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID-1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE-98295987; PubMed-9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmier K., Gass S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Davis K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagals K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RL laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CBHY/CBBZ/GPH/YIEH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z70692; CA94666.1; ALT_TERM.
 CC EMBL; Z70692; CA94655.1; ALT_INIT.
 CC TIGR; AE007074; AAK46576.1; -
 CC Tuberculist; RV2232; -
 CC Tuberculist; RV2233; -
 CC InterPro; IPR005834; Hydrolase.

DR Pfam; PF00702; Hydrolase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 291 AA; 30694 MW; 750F090FB154E5E5 CRC64;
 Query Match 43.8%; Score 7; DB 1; Length 291;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RAARAA 12
 DB 66 RAARAA 72

RESULT 3
 PROA_MEIRU STANDARD; PRT; 417 AA.
 ID PROA_MEIRU
 AC O86053;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gamma-glutamyl phosphate reductase (GPR) (EC 1.2.1.41) (Glutamate-5-
 DE semialdehyde dehydrogenase) (Glutamyl-gamma-semialdehyde
 DE dehydrogenase) (GSA dehydrogenase).
 GN PROA.
 OS Meiothermus ruber.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 CC Meiothermus.
 OX NCBI_TaxID-277;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-40;
 RA Yaklichkin S.Y., Zimina M.S., Yurchenko Y.V., Hromov I.S.,
 RA Neumivakin L.V.;
 RT "Molecular cloning and sequence analysis of the proA gene from
 RT thermophilic eubacterium Thermus ruber."
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE NADPH DEPENDENT REDUCTION OF L-GAMMA-
 CC GLUTAMYL 5-PHOSPHATE INTO L-GLUTAMATE 5-SEMIALDEHYDE AND
 CC PHOSPHATE. THE PRODUCT SPONTANEOUSLY UNDERGOES CYCLIZATION TO FORM
 CC 1-PYRROLINE-5-CARBOXYLATE.
 CC -1- CATALYTIC ACTIVITY: L-glutamate 5-semialdehyde + phosphate +
 CC NADP(+) = L-gamma-glutamyl 5-phosphate + NADPH.
 CC -1- PATHWAY: Proline biosynthesis; second step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE GAMMA-GLUTAMYL PHOSPHATE REDUCTASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF082661; AAC72811.1; -
 CC HAMAP; MF 00412; -; 1.
 CC InterPro; IPR002086; Aldehyde_dehydr.
 CC InterPro; IPR000965; Gglut_pp_reduct.
 CC Pfam; PF00171; aldehyd; 1.
 CC TIGR; TIGR00407; proA; 1.
 CC PROSITE; PS01223; proA; 1.
 CC Oxidoreductase; Proline biosynthesis; NADP.
 CC SEQUENCE 417 AA; 44985 MW; A280A8A7E9C92268 CRC64;
 Query Match 43.8%; Score 7; DB 1; Length 417;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ARAAAR 10
 DB 13 ARAAAR 19

RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus";
RL Nature 409:1007-1011(2001).
CC -1- SIMILARITY: TO M.TUBERCULOSIS RV0486.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U00018; AAA17228.1; ALT_INT.
DR EMBL; AL563925; CAC31960.1; -
DR PIR; H87214; H87214.
DR Leproma; ML2443; -
DR InterPro; IPR001296; Glyco_trans.1.
DR Pfam; PF00534; Glyco_transf.1.1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 428 AA; 45391 MW; A149F0187E3587C CRC64;

Query Match 43.8%; Score 7; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRAARA 8
DB 204 RRAARA 210
|||||||

RESULT 6
VGLX_HSVS STANDARD; PRT; 444 AA.
ID VGLX_HSVS
AC Q08103;
DR 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Glycoprotein GX precursor.
OS Bovine herpesvirus type 1.2 (strain ST).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
ON NCBI_TaxID=45407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-94167875; PubMed-8122370;
RA Leung-Tack P., Audonnet J.P., Riviere M.;
RT "The complete DNA sequence and the genetic organization of the short
RT unique region (US) of the bovine herpesvirus type 1 (ST strain).";
RL Virology 199:409-421(1994).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z23068; CAAB0603.1; -
DR PIR; S35783; S35783.
DR InterPro; IPR003363; Herpes_gg.
DR Pfam; PF02400; Herpes_gg.1.
KW Glycoprotein; Transmembrane; Signal.
FT CHAIN 1 24 POTENTIAL.
FT SIGNAL 25 44 GLYCOPROTEIN GX..
FT TRANSME 390 414 POTENTIAL.
FT CARBOHYD 117 117 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 444 AA; 46708 MW; 0145942AA35B05CB CRC64;

Query Match 43.8%; Score 7; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AARRARA 14
DB 422 AARRARA 428
|||||||

RESULT 7
CYAE_BORPE STANDARD; PRT; 474 AA.
ID CYAE_BORPE
AC P11092;
DR 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Protein cyae precursor.
GN CYAE.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=18323;
RX MEDLINE-89091151; PubMed-2905265;
RA Glaser P., Sakamoto H., Bellalou J., Ullmann A., Danchin A.;
RT "Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-
RT haemolysin bifunctional protein of Bordetella pertussis.";
RL EMBL J. 7:3997-4004(1988).
CC -1- FUNCTION: CYAE IS NECESSARY FOR TRANSPORT OF CALMODULIN-SENSITIVE
CC ADENYLATE CYCLASE-HEMOLYSIN (CYCLOLYSIN).
CC -1- SUBCELLULAR LOCATION: Outer membrane (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PRTF FAMILY OF SECRETION PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X14199; CAA32414.1; -
DR PIR; S02388; BVBCE.
DR InterPro; IPR003423; OEP.
DR Pfam; PF02321; OEP; 2.
KW Hemolysis; Transport; Outer membrane; Signal.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 474 PROTEIN CYAE.
SQ SEQUENCE 474 AA; 50204 MW; 29A4F21D377FC957 CRC64;

Query Match 43.8%; Score 7; DB 1; Length 474;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AARRARA 14
DB 203 AARRARA 209
|||||||

RESULT 8
Y486_MYCTU STANDARD; PRT; 480 AA.
ID Y486_MYCTU
AC Q11152;
DR 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Hypothetical protein RV0486.
GN RV0486 OR MT0504 OR MTCY2069.12.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteridae; Actinomycetales;

```

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO M.LSPRAE ML2443.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 277162; CAB00947.1;
DR EMBL; AE006951; AAK44727.1;
DR PIR; A70744; A70744.
DR TIGR; MT0504;
DR TubercuList; RV0486;
DR InterPro; IPR001296; Glyco_trans_1.
DR Pfam; PF00534; Glycos_transf_1;
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 480 AA; 50541 MW; 2134755E894A9CCF CRC64;
SQ
Query Match 43.88; Score 7; DB 1; Length 480;
Best Local Similarity 100.08; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2 RRAAARA 8
Db 252 RRAAARA 258
[1]
[2]
[3]
[4]
[5]
[6]
[7]
[8]
[9]
[10]
[11]
[12]
[13]
[14]
[15]
[16]
[17]
[18]
[19]
[20]
[21]
[22]
[23]
[24]
[25]
[26]
[27]
[28]
[29]
[30]
[31]
[32]
[33]
[34]
[35]
[36]
[37]
[38]
[39]
[40]
[41]
[42]
[43]
[44]
[45]
[46]
[47]
[48]
[49]
[50]
[51]
[52]
[53]
[54]
[55]
[56]
[57]
[58]
[59]
[60]
[61]
[62]
[63]
[64]
[65]
[66]
[67]
[68]
[69]
[70]
[71]
[72]
[73]
[74]
[75]
[76]
[77]
[78]
[79]
[80]
[81]
[82]
[83]
[84]
[85]
[86]
[87]
[88]
[89]
[90]
[91]
[92]
[93]
[94]
[95]
[96]
[97]
[98]
[99]
[100]
[101]
[102]
[103]
[104]
[105]
[106]
[107]
[108]
[109]
[110]
[111]
[112]
[113]
[114]
[115]
[116]
[117]
[118]
[119]
[120]
[121]
[122]
[123]
[124]
[125]
[126]
[127]
[128]
[129]
[130]
[131]
[132]
[133]
[134]
[135]
[136]
[137]
[138]
[139]
[140]
[141]
[142]
[143]
[144]
[145]
[146]
[147]
[148]
[149]
[150]
[151]
[152]
[153]
[154]
[155]
[156]
[157]
[158]
[159]
[160]
[161]
[162]
[163]
[164]
[165]
[166]
[167]
[168]
[169]
[170]
[171]
[172]
[173]
[174]
[175]
[176]
[177]
[178]
[179]
[180]
[181]
[182]
[183]
[184]
[185]
[186]
[187]
[188]
[189]
[190]
[191]
[192]
[193]
[194]
[195]
[196]
[197]
[198]
[199]
[200]
[201]
[202]
[203]
[204]
[205]
[206]
[207]
[208]
[209]
[210]
[211]
[212]
[213]
[214]
[215]
[216]
[217]
[218]
[219]
[220]
[221]
[222]
[223]
[224]
[225]
[226]
[227]
[228]
[229]
[230]
[231]
[232]
[233]
[234]
[235]
[236]
[237]
[238]
[239]
[240]
[241]
[242]
[243]
[244]
[245]
[246]
[247]
[248]
[249]
[250]
[251]
[252]
[253]
[254]
[255]
[256]
[257]
[258]
[259]
[260]
[261]
[262]
[263]
[264]
[265]
[266]
[267]
[268]
[269]
[270]
[271]
[272]
[273]
[274]
[275]
[276]
[277]
[278]
[279]
[280]
[281]
[282]
[283]
[284]
[285]
[286]
[287]
[288]
[289]
[290]
[291]
[292]
[293]
[294]
[295]
[296]
[297]
[298]
[299]
[300]
[301]
[302]
[303]
[304]
[305]
[306]
[307]
[308]
[309]
[310]
[311]
[312]
[313]
[314]
[315]
[316]
[317]
[318]
[319]
[320]
[321]
[322]
[323]
[324]
[325]
[326]
[327]
[328]
[329]
[330]
[331]
[332]
[333]
[334]
[335]
[336]
[337]
[338]
[339]
[340]
[341]
[342]
[343]
[344]
[345]
[346]
[347]
[348]
[349]
[350]
[351]
[352]
[353]
[354]
[355]
[356]
[357]
[358]
[359]
[360]
[361]
[362]
[363]
[364]
[365]
[366]
[367]
[368]
[369]
[370]
[371]
[372]
[373]
[374]
[375]
[376]
[377]
[378]
[379]
[380]
[381]
[382]
[383]
[384]
[385]
[386]
[387]
[388]
[389]
[390]
[391]
[392]
[393]
[394]
[395]
[396]
[397]
[398]
[399]
[400]
[401]
[402]
[403]
[404]
[405]
[406]
[407]
[408]
[409]
[410]
[411]
[412]
[413]
[414]
[415]
[416]
[417]
[418]
[419]
[420]
[421]
[422]
[423]
[424]
[425]
[426]
[427]
[428]
[429]
[430]
[431]
[432]
[433]
[434]
[435]
[436]
[437]
[438]
[439]
[440]
[441]
[442]
[443]
[444]
[445]
[446]
[447]
[448]
[449]
[450]
[451]
[452]
[453]
[454]
[455]
[456]
[457]
[458]
[459]
[460]
[461]
[462]
[463]
[464]
[465]
[466]
[467]
[468]
[469]
[470]
[471]
[472]
[473]
[474]
[475]
[476]
[477]
[478]
[479]
[480]
[481]
[482]
[483]
[484]
[485]
[486]
[487]
[488]
[489]
[490]
[491]
[492]
[493]
[494]
[495]
[496]
[497]
[498]
[499]
[500]
[501]
[502]
[503]
[504]
[505]
[506]
[507]
[508]
[509]
[510]
[511]
[512]
[513]
[514]
[515]
[516]
[517]
[518]
[519]
[520]
[521]
[522]
[523]
[524]
[525]
[526]
[527]
[528]
[529]
[530]
[531]
[532]
[533]
[534]
[535]
[536]
[537]
[538]
[539]
[540]
[541]
[542]
[543]
[544]
[545]
[546]
[547]
[548]
[549]
[550]
[551]
[552]
[553]
[554]
[555]
[556]
[557]
[558]
[559]
[560]
[561]
[562]
[563]
[564]
[565]
[566]
[567]
[568]
[569]
[570]
[571]
[572]
[573]
[574]
[575]
[576]
[577]
[578]
[579]
[580]
[581]
[582]
[583]
[584]
[585]
[586]
[587]
[588]
[589]
[590]
[591]
[592]
[593]
[594]
[595]
[596]
[597]
[598]
[599]
[600]
[601]
[602]
[603]
[604]
[605]
[606]
[607]
[608]
[609]
[610]
[611]
[612]
[613]
[614]
[615]
[616]
[617]
[618]
[619]
[620]
[621]
[622]
[623]
[624]
[625]
[626]
[627]
[628]
[629]
[630]
[631]
[632]
[633]
[634]
[635]
[636]
[637]
[638]
[639]
[640]
[641]
[642]
[643]
[644]
[645]
[646]
[647]
[648]
[649]
[650]
[651]
[652]
[653]
[654]
[655]
[656]
[657]
[658]
[659]
[660]
[661]
[662]
[663]
[664]
[665]
[666]
[667]
[668]
[669]
[670]
[671]
[672]
[673]
[674]
[675]
[676]
[677]
[678]
[679]
[680]
[681]
[682]
[683]
[684]
[685]
[686]
[687]
[688]
[689]
[690]
[691]
[692]
[693]
[694]
[695]
[696]
[697]
[698]
[699]
[700]
[701]
[702]
[703]
[704]
[705]
[706]
[707]
[708]
[709]
[710]
[711]
[712]
[713]
[714]
[715]
[716]
[717]
[718]
[719]
[720]
[721]
[722]
[723]
[724]
[725]
[726]
[727]
[728]
[729]
[730]
[731]
[732]
[733]
[734]
[735]
[736]
[737]
[738]
[739]
[740]
[741]
[742]
[743]
[744]
[745]
[746]
[747]
[748]
[749]
[750]
[751]
[752]
[753]
[754]
[755]
[756]
[757]
[758]
[759]
[760]
[761]
[762]
[763]
[764]
[765]
[766]
[767]
[768]
[769]
[770]
[771]
[772]
[773]
[774]
[775]
[776]
[777]
[778]
[779]
[780]
[781]
[782]
[783]
[784]
[785]
[786]
[787]
[788]
[789]
[790]
[791]
[792]
[793]
[794]
[795]
[796]
[797]
[798]
[799]
[800]
[801]
[802]
[803]
[804]
[805]
[806]
[807]
[808]
[809]
[810]
[811]
[812]
[813]
[814]
[815]
[816]
[817]
[818]
[819]
[820]
[821]
[822]
[823]
[824]
[825]
[826]
[827]
[828]
[829]
[830]
[831]
[832]
[833]
[834]
[835]
[836]
[837]
[838]
[839]
[840]
[841]
[842]
[843]
[844]
[845]
[846]
[847]
[848]
[849]
[850]
[851]
[852]
[853]
[854]
[855]
[856]
[857]
[858]
[859]
[860]
[861]
[862]
[863]
[864]
[865]
[866]
[867]
[868]
[869]
[870]
[871]
[872]
[873]
[874]
[875]
[876]
[877]
[878]
[879]
[880]
[881]
[882]
[883]
[884]
[885]
[886]
[887]
[888]
[889]
[890]
[891]
[892]
[893]
[894]
[895]
[896]
[897]
[898]
[899]
[900]
[901]
[902]
[903]
[904]
[905]
[906]
[907]
[908]
[909]
[910]
[911]
[912]
[913]
[914]
[915]
[916]
[917]
[918]
[919]
[920]
[921]
[922]
[923]
[924]
[925]
[926]
[927]
[928]
[929]
[930]
[931]
[932]
[933]
[934]
[935]
[936]
[937]
[938]
[939]
[940]
[941]
[942]
[943]
[944]
[945]
[946]
[947]
[948]
[949]
[950]
[951]
[952]
[953]
[954]
[955]
[956]
[957]
[958]
[959]
[960]
[961]
[962]
[963]
[964]
[965]
[966]
[967]
[968]
[969]
[970]
[971]
[972]
[973]
[974]
[975]
[976]
[977]
[978]
[979]
[980]
[981]
[982]
[983]
[984]
[985]
[986]
[987]
[988]
[989]
[990]
[991]
[992]
[993]
[994]
[995]
[996]
[997]
[998]
[999]
[1000]

```

Sat Aug 9 19:18:31 2003

*Analysis of the protein-coding content of the sequence of human

RT Cytomegalovirus strain AD169. 154:125-169(1990).
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC -1- SIMILARITY: BELONGS TO THE US22 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL: X17403; CAA35293.1; -
 CC EMBL: X04650; CAB37117.1; -
 CC PIR: A27216; QOBED1.
 CC InterPro: IPR003360; US22.
 CC Pfam: PF02393; US22; 1.
 CC Hypothetical protein.
 CC CARBOHYD 44 N-LINKED (GLCNAC...) (POTENTIAL).
 CC SEQUENCE 603 AA; 70020 MW; E34F64D01E27687C CRC64;
 CC

Query Match 43.8%; Score 7; DB 1; Length 603;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRAARAA 8

DB 574 RRAARAA 580

RESULT 11.

HEMLAGABI
 ID HEMIAGABI STANDARD; PRT; 621 AA.
 AC Q92403;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE 5-aminolevulinic acid synthase, mitochondrial precursor (EC 2.3.1.37)
 DE (Delta-aminolevulinic synthase) (Delta-ALA synthetase).
 GN HEMI.
 OS Agaricus bisporus (Common mushroom).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Agaricales; Agaricaceae; Agaricus.
 CC NCBI_TaxID=5341;
 CC SEQUENCE FROM N.A.
 CC [1]
 CC STRAIN=D649;
 CC Yague E., Mehak-Zunic M., Wood D.A., Thurston C.F.;
 CC Submitted (NOV-1996) to the EMBL/GenBank/DDJJ databases.
 CC -1- CATALYTIC ACTIVITY: Succinyl-CoA + glycine -> 5-aminolevulinic acid +
 CC CoA + CO(2).
 CC -1- COFACTOR: Pyridoxal phosphate.
 CC -1- PATHWAY: Heme biosynthesis; first (rate-limiting) step.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- SIMILARITY: BELONGS TO CLASS-II OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL: Z50096; CAA90424.1; -
 CC InterPro: IPR003408; Ala_synthase.
 CC InterPro: IPR004839; Aminotransf1/2.
 CC InterPro: IPR001917; Ntrtransf_2.
 CC Pfam: PF02490; Ala_synthase; 1.
 CC Pfam: PF00155; aminotran_1_2; 1.

RESULT 13

DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
 KW Heme biosynthesis; Transferase; Acyltransferase; Mitochondrion;
 KW Transit peptide; Pyridoxal phosphate.
 FT TRANSIT 1 ? MITOCHONDRION.
 FT CHAIN 1 ? 621 5-AMINOLEVULINIC ACID SYNTHASE.
 FT BINDING 362 362 PYRIDOXAL PHOSPHATE (PROBABLE).
 FT SEQUENCE 621 AA; 67426 MW; A335C3268FAE1AA3 CRC64;
 SQ

Query Match 43.8%; Score 7; DB 1; Length 621;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ARAAAR 10

DB 54 ARAAAR 60

RESULT 12

UL47_HSVBP STANDARD; PRT; 742 AA.

ID UL47_HSVBP
 AC P30021; 1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 29, Last annotation update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE 80.7 kDa alpha trans-inducing protein (VP8 tegument protein) (VP7
 DE protein) (107 kDa protein).
 DE VP8.
 GN Bovine herpesvirus type 1 (strain P8-2).
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 CC NCBI_TaxID=10324;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=92113550; PubMed=1663698;
 CC Carpenter D.E., Misra V.;
 CC "The most abundant protein in bovine herpes 1 virions is a homologue
 CC of herpes simplex virus type 1 UL47.";
 CC J. Gen. Virol. 72:3077-3084(1991).
 CC -1- FUNCTION: MODULATOR OF ALPHA-TIF (VNM65 PHOSPHOPROTEIN) TRANS-
 CC ACTIVATION. UL47 MAY HAVE KINASE ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: MAJOR TEGUMENT PROTEIN OF THE VIRIONS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE LATER STAGES OF INFECTION.
 CC -1- PTM: PHOSPHORYLATED.
 CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL47,
 CC EHV-1 13, AND VZV 11.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL: D10327; BAA01170.1; -
 CC EMBL: Z11610; CAA77683.1; -
 CC PIR: JQ1435; TNBE1.
 CC InterPro: IPR005029; Herpes_UL47.
 CC Pfam: PF03362; Herpes_UL47; 1.
 CC Transcription regulation; Trans-acting factor; Structural protein;
 CC Late protein; Phosphorylation.
 CC SEQUENCE 742 AA; 80744 MW; 85979D8C2C953C89 CRC64;
 SQ

Query Match 43.8%; Score 7; DB 1; Length 742;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ARAAAR 10

DB 376 ARAAAR 382


```

RESULT 15
VP14_EBV
ID VP14_EBV STANDARD; PRT; 1318 AA.
AC P03179;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Probable membrane antigen P140 (tegument protein).
GN BNRF1.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10377;
RN [1]
RP COMPLETE GENOME.
RX MEDLINE=84270667; PubMed=60871149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tufnell P.S., Barrell B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86045995; PubMed=2998073;
RA Hudson G.S., Bankier A.T., Satchwell S.C., Barrell B.G.;
RT "The short unique region of the B95-8 Epstein-Barr virus genome.";
RL Virology 147:81-98(1985).
CC -!- FUNCTION: TEGUMENT PROTEIN.
CC -!- SIMILARITY: TO BOTH 140 kDa MEMBRANE ANTIGEN ENCODED AT THE
CC EXTREMITIES OF THE HERPESVIRUS SAIMIRI GENOME (3 AND 75/EILF1).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01555; CAA24862.1; -
DR EMBL; M11924; AAA45899.1; -
DR PIR; A03740; QDBE1.
DR InterPro; IPR000728; AIRS_related.
DR Pfam; PF02769; AIRS_C; 1.
KW Membrane; Antigen; Late protein; Structural protein.
SQ SEQUENCE 1318 AA; 142843 MW; 58D1DC644EB4BE6 CRC64;

Query Match 43.8%; Score 7; DB 1; Length 1318;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ARAAARR 11
   |||||
Db 322 ARAAARR 328

Search completed: August 9, 2003, 16:29:52
Job time : 9.22857 secs

```